

# Start-up Vibration Analysis for Novelty Detection on Industrial Gas Turbines

Yu Zhang, Samuel Cruz-Manzo

School of Engineering  
University of Lincoln  
Lincoln, U.K.

{yzhang, scruzmanzo}@lincoln.ac.uk

Anthony Latimer

Siemens Industrial Turbomachinery LTD  
Lincoln, U.K.

[anthony.latimer@siemens.com](mailto:anthony.latimer@siemens.com)

**Abstract**— This paper focuses on industrial application of start-up vibration signature analysis for novelty detection with experimental trials on industrial gas turbines (IGTs). Firstly, a representative vibration signature is extracted from healthy start-up vibration measurements through the use of an adaptive neuro-fuzzy inference system (ANFIS). Then, the first critical speed and the vibration level at the critical speed are located from the signature. Finally, two (s- and v-) health indices are introduced to detect and identify different novel/fault conditions from the IGT start-ups, in addition to traditional similarity measures, such as Euclidean distance and cross-correlation measures. Through a case study on IGTs, it is shown that the presented approach provides a convenient and efficient tool for IGT condition monitoring using start-up field data.

**Keywords**— *Start-up vibration signature; adaptive neuro-fuzzy inference system; s- & v- health indices; novelty/fault detection; industrial gas turbine.*

## I. INTRODUCTION

The use of industrial gas turbines (IGTs) is widespread. The considered IGT units use direct or indirect-coupled power turbines that develop energy to supply equipment such as compressors, pumps or generators. An essential part of these complex systems are the control and monitoring units, which play an important role in ensuring system safety and performance [1].

Vibration, especially in fluid film bearings for instance, as used in IGTs, can have an extremely detrimental effect on lifetime and performance, and contribute to significant repair costs and lost production [2]. Vibration measurements in steady states, especially with a low sampling rate for IGTs, can be useful for fault detection after bearing or shaft faults have occurred [3]. Alternatively, measurements from non-steady states, for instance start-up vibration data, contain rich information regarding the health condition of the machines. This is due to the fact that certain fault conditions may result in a change in its rotational dynamics, which results in a change of resonant peak in the vibration signature [4]. Vibration signature analysis for start-ups can provide early warning of potential faults, such as shaft imbalance, blade and shaft crack propagation and other equipment malfunctions.

This paper is of interest for a wide engineering society. It considers the problem of start-up vibrations, which has been solved by introducing the threshold in vibration measurements

for decades. The aim of this paper is (1) to search for a representative vibration signature from healthy start-up signals, and (2) to introduce health indices to detect and identify different novelty/fault conditions within the gas turbine.

For IGT systems, pure dynamic models, which capture detailed physical properties of the start-up operation, are normally difficult to obtain due to complexity of structural coupling, auxiliaries and control system. By contrast, the use of black-box models, such as artificial neural networks (ANNs), has therefore been utilised widely because of their non-linear and non-parametric properties [5]. For instance, [6] has presented the use of ANNs in condition motoring of IGTs for performance trend change detection. Gas turbine health monitoring has been carried on in [7] using ANNs for a high bypass ratio military turbofan engine, where double-component faults have been presented to demonstrate the capability of ANNs as a diagnostic tool. Applications of ANNs have shown some success, however, the black-box model provides little or no insight of the physical underlying attributes of such complex systems, and as a result is not effective with uncertainties [8].

On the other hand, fuzzy systems can manipulate and interpret relationship between system inputs and outputs by fuzzy rules, and then is more effective in dealing with uncertainties. Therefore, a hybridization of ANN and fuzzy systems is desirable, for example an adaptive neuro-fuzzy inference system (ANFIS) [9]. ANFIS corresponds to a set of fuzzy rules, and in each rule there is a neural network for parameter estimation. This combination captures benefits from both fuzzy rules –e.g. expert knowledge on known distributions of inputs or outputs– and ANN for non-linear non-parametric estimation. ANFIS has been used extensively since its invention, for instance, [10] has developed a data-driven fault detection methodology for an industrial steam turbine using three ANFIS classifiers fed with the most influential diagnostic information. A model-based fault diagnosis system has been created in [11] on a gas turbine power plant by applying ANFIS to approximate the non-linear Rowen's model. [12] has also investigated condition monitoring and fault diagnosis in Francis turbine based on integration of numerical modelling of turbine runner with multiple ANNs and multiple ANFIS techniques.

In this paper, ANFIS is applied to extract the vibration signature from 'healthy' start-up data through an offline

training model. Subsequently, on an on-line testing platform, a new input start-up signal from a running turbine is fed in to compare with the extracted vibration signature, and finally anomalies or faults can be identified by investigation of the introduced similarity measures.

Traditionally, similarity measures between two signals can be obtained through Euclidean distance and cross-correlation [13,14]. However, through practice from this paper, it is proven that the introduced s- (speed-related) and v- (vibration-related) health indices are more appropriate choices for this industrial application. For a sub-15MW IGT considered here, start-up measurements from experimental trials are used to demonstrate the proposed approach to be practical and useful.

## II. METHODOLOGY

The use of ANN for IGT Rotor Shaft Diagnosis has been reported by [15]. A mathematical model for vibration analysis was required to train ANN and capture vibration-speed dependency. This represents a challenge as it is necessary to take into account the interaction of conservation of mechanical energy across the mechanical system as well as losses in mechanical components. In this paper, ANFIS to extract the vibration signature from ‘healthy’ start-up data through an offline training model is considered as a first approach for vibration analysis on IGTs. In a future work, a comparison between ANN, ANFIS and fuzzy inference system (FIS) for Start-up Vibration Analysis in IGTs will be investigated.

The methodology is depicted in Table 1 for brevity, with an overview of underpinning principles described in the following sub-sections.

It is worthwhile mentioning that the ‘offline training’ (e.g. with complex physical and numerical models) and ‘on-line testing’ (e.g. with simple mathematical expressions) strategy gains particular favour in terms of implementation in an industrial environment.

TABLE I. OUTLINE METHODOLOGY

Steps	Inputs	Methods	Results
Data collection	Running turbine during startup	Machine probe sensors	Vibration and speed measurements
Offline training	Vibration vs. speed data during startups	ANFIS	Startup vibration signature
On-line testing	New startup vibration vs. speed signal	Comparing similarity measures	Health indices
Novelty detection	Health indices	Statistical analysis	Warning on anomalies/faults in start-ups

### A. ANFIS

ANFIS is a powerful signal/information processing tool which combines fuzzy logic and ANN [9]. ANFIS uses fuzzy rules, where ANN –normally with a BP learning algorithm– is used for membership function approximation from input and output data. For further reading, [16] has given a detailed explanation of the principles of FIS and ANN with BP.

A typical structure of ANFIS is shown in Fig. 1, with one input (e.g. speed), one output (e.g. vibration) and 5 rules [17]. This example is based on Training set 1 data in Section 3 just to provide a seed of demonstration on the ANFIS. Here, each membership function of the input (Gaussian function in this case) corresponds to a linear membership function of the output. For instance, when input = 60 (percentage of full speed), Rule 3’s output =  $0.4481 \times 60 - 3.682$ , and the same applies to the other rules, then at the end the final output is calculated by a weighted sum of the results from all rules, as shown in Fig. 1.

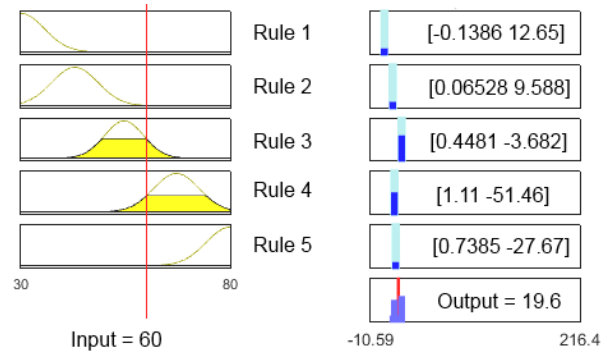


Fig. 1. System ANFIS: input and output rules

### B. Critical speed and health indices

During a start-up or shutdown, the engine will normally pass through at least one critical rotor speed or natural frequency. This condition occurs when the forcing frequency of the rotor matches that of one of the natural modes of vibration of the system, i.e. transverse, torsional etc. This leads to a resonant peak in the vibration signature [4].

For matching a new signal  $\mathbf{a}$  with the extracted ‘healthy’ vibration signature  $\mathbf{b}$ , the similarity measures are defined as

*Euclidean distance* =  $\|\mathbf{a} - \mathbf{b}\|$ , and

*Angle between vectors*  $\cos \alpha = \frac{\mathbf{a} \cdot \mathbf{b}}{\|\mathbf{a}\| \|\mathbf{b}\|}$

The measure of angle between vectors (ABV) is related to the Cross-Correlation measure. Here, ABV is used to avoid confusions. Generally, these two measures – Euclidean distance and ABV measure – could produce sufficient information on the similarities between two signals, where Euclidean distance measures the distances/area between two signals, while the ABV measures the rotation of two signals at a defined point of interest.

For the identified peak point  $(s_a, v_a)$  in the new signal  $a$  and the peak point  $(s_b, v_b)$  in the extracted signature  $b$ , two (s- and v-) health indices are defined as

$$s\text{-health index} = \left(1 - \frac{|s_a - s_b|}{|s_b|}\right) \times 100, \text{ and}$$

$$v\text{-health index} = \left(1 - \frac{v_a - v_b}{|v_b|}\right) \times 100.$$

This implies that the indices of 100 indicate 100% healthy condition, and the lower the indices, the more differences between the speed or vibration with that of the ‘healthy’ vibration signature.

It is noticed that we are using absolute difference of speed values and actual difference of vibration values. It means that the v-index can go over 100, indicating an over-healthy situation with lower vibrations during start-ups, in which case traditional Euclidean distances may produce a false alarm by measuring the distances between two signals.

For further comparison, the Euclidean distance and ABV measures are modified to the percentage health format, as follows

$$\text{Corrected Euclidean distance} = \left(1 - \frac{\|a - b\|}{\|b\|}\right) \times 100, \text{ and}$$

$$\text{Corrected angle between vectors} = \frac{a \cdot b}{\|a\| \|b\|} \times 100.$$

In this case, s-health index is comparable to corrected ABV, and v-health index to corrected Euclidean distance.

### III. CASE STUDY

#### A. Experimental data measurements

Vibration and speed measurements from a sub-15MW IGT are utilised here for case study. In this paper, the vibration measurement is recorded from an IGT inlet bearing, and the speed measurement indicates the rotor speed (in terms of percentage of the full speed). Sample speed and vibration signals in time domain and the corresponding vibration signatures are shown in Fig. 2 separately.

Due to the noise nature of the data sets for the starting and ending sections, only the important features to demonstrate rotor dynamics at the first critical speed—here within 30—80% of full speed—are used. In the following sections, 5 speed vs. vibration data sets are selected as training data, which are considered as ‘healthy’ start-ups by industrial experts; and 5 more data sets, including normal and abnormal data, are used for testing purpose.

#### B. Offline training: start-up signature extraction

With information from 5 training sets, a start-up vibration signature, as shown in Fig. 3, is extracted through the use of an ANFIS with speed data as input and vibration data as output. It is apparent from the extracted signature that a peak (first critical speed) and a valley are present.

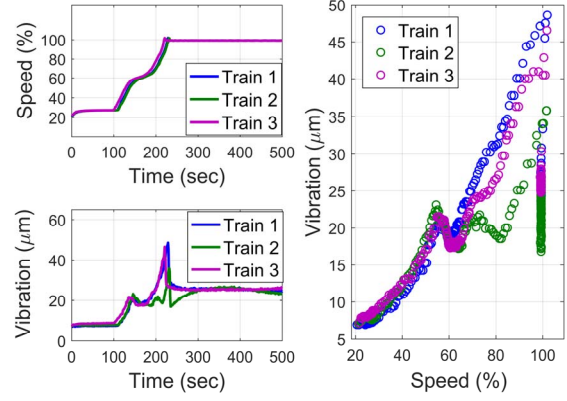


Fig. 2. Example start-up signals: speed measurements in time domain, vibration measurements in time domain and vibration signatures (Train = Examples, Training data for next subsection)

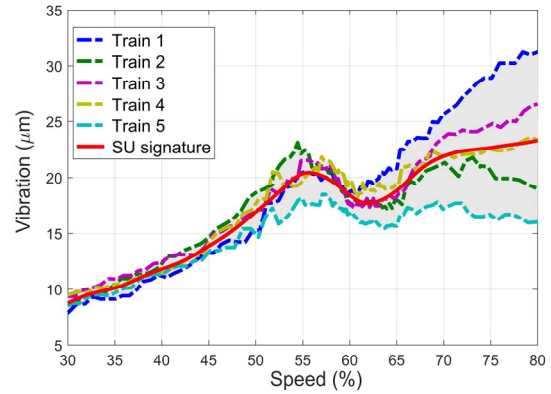


Fig. 3. Start-up signature extraction by ANFIS (Train = Training data; SU signature = extracted Start-Up vibration signature)

#### C. On-line testing: critical speed and matching

Here, 5 testing data sets are utilised for assessment, as shown in Fig. 4. Test data sets 1 and 2 are from ‘normal’ operations of the same type of engine (i.e. different engine but with same assembly configurations); Test data sets 3 and 4 are from the same engine as Training sets but with excessive ‘abnormal’ vibrations; and Test data set 5 is from testing set of a different type of engine (different core configurations).

A smoothing technique, again by ANFIS, is used to each testing signal in order to present peak and valley clearly. Then local maxima and minima can be found using ‘findpeak’ function in Matlab® [17]. The valley may be related to some health or fault information, although in this paper, only the peak point is considered. Finally, the first critical speed and corresponding vibration, i.e. the peak point, is compared with that of the extracted ‘healthy’ vibration signature, and the s- and v-health indices are calculated.

Here, ANFIS and ‘findpeak’ are applied as a proof of concept, whereas for implementation in real industrial environment, different techniques with lower computational complexity will be used.

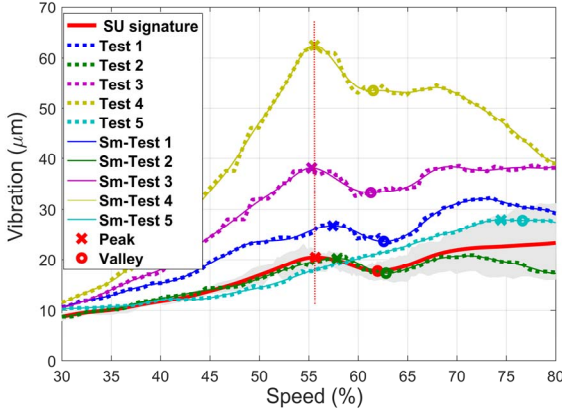


Fig. 4. Start-up testing (Test = Testing data; Sm-Test = Smoothened Testing data; X indicates the peak point and O indicates the valley point in all cases)

#### D. Novelty detection

It is noticeable that different engine configurations (core change) can cause changes in critical speeds. Moreover, severe vibrations can ruin bearings, damage shafts and potentially cause impairment of the whole machine. For both cases to be identified, the s- and v- health indices, with s-index indicating core changes, and v-index indicating excessive vibration levels, have therefore provided essential information for novelty detection on different potential fault situations.

The similarity measures, traditional Euclidean distance and ABV, are converted to corrected health indices in percentage form for comparison purpose. The comparisons of 4 indices are plotted in Fig. 5. Considering 100 being 100% healthy, in this case, the lower threshold is set as the lowest value obtained from the 5 train sets, and the Pass/Fail results of the 5 testing data sets are shown in Table II.

From the results, it is shown that s-index is useful to identify core changes, which is more advantageous over ABV measures, whilst v-index is useful to identify excessive vibrations, which gives comparable results to Euclidean distance measure. The ABV and Euclidean distance measures could not identify the condition of core change in Test 5, which s- and v- indices have identified correctly. On the opposite, for Test 4, although there are excessive vibrations, s-index can indicate that there is no core change, which may be important information for further fault diagnosis. It is also shown that, there are insignificant differences of the health indices between different engines with the same configurations (Test 1 & 2 with the Training sets), which may indicate that for the same type of engines, the representative vibration signature could be extracted and compared altogether.

#### E. Discussions

This case study works as a proof of concept on start-up vibration signature matching for novelty/fault detection of IGTs. Here, limitations at the current stage and future recommendations are listed:

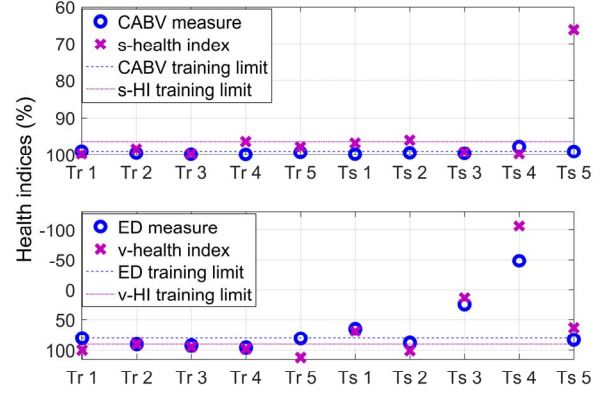


Fig. 5. Comparison of similarity measures/health indices (CABV = corrected ABV; ED = corrected Euclidean Distance; HI = Health Index; Tr = Training data set; Ts = Testing data set)

- *Offline training:* More ‘healthy’ training sets are required to extract a more representative start-up vibration signature of the engine.
- *Signature extraction:* In terms of choosing methodology for signature extraction, e.g. ANFIS, ANN or FIS, maybe the question to be answered is: which one is able to assure the obtained vibration(speed) dependency really has the essential peak/value characteristics of the healthy start-up, and the peak and valley information is properly ‘averaged’ over the training healthy start-ups. This is yet to be investigated.

On the other hand, the technique could be improved, where optimal non-uniform segments of the input can be obtained using an optimization algorithm. In this way, the regions of interest or of less interest can be revealed. A piece of preliminary work is conducted, as shown in Fig. 6, where for the testing data, the number of frames/rules is determined as six through the use of variational Bayesian inference, and the associated (Gaussian) membership functions are calculated which correspond to the six rules. It can be seen that Rule 3 corresponds to the valley region, while Rule 4 corresponds to the peak region in the signature. These rules can recognisably provide more meaningful information than the original ANFIS system as was shown in Fig. 1.

- *Similarity measures:* Following the last discussion point, regional similarity measures –measures from each non-uniform segment– can be introduced, which may provide more detailed health or fault information.
- *Industrial implementation:* Considering the application on industrial data and software platform, offline training strategy is a suitable choice to avoid complexity. For online testing purpose, taking into account of industrial environment, a simple smoothing technique, e.g. moving average, and a simple peak finding technique, e.g. with gradient searching, are implemented in C code and trailed.

TABLE II. COMPARISON OF SIMILARITY MEASURES/ HEALTH INDICES

	Train 1	Train 2	Train 3	Train 4	Train 5	Test 1	Test 2	Test 3	Test 4	Test 5
<i>C.ABV.</i>	99.09	99.58	99.89	99.94	99.40	99.88	99.52	99.66	97.82	99.18
	T.L.	—	—	—	—	V	V	V	X	V
<i>s-HI</i>	99.85	98.45	99.79	96.42	97.86	96.83	96.05	99.33	99.82	66.21
	—	—	—	T.L.	—	V	X	V	V	X
<i>E.D.</i>	80.71	90.69	92.79	96.12	81.10	65.31	88.53	25.21	-48.03	83.33
	T.L.	—	—	—	—	X	V	X	X	V
<i>v-HI</i>	100.26	90.65	95.00	98.24	112.13	69.32	101.10	13.53	-106.10	63.45
	—	T.L.	—	—	—	X	V	X	X	X

(C.ABV. = corrected ABV; C.E.D. = corrected Euclidean Distance; HI = health index; Train = Training data set; Test = Testing data set; T.L. = Training Limit; V = 'pass'; X = 'fail')

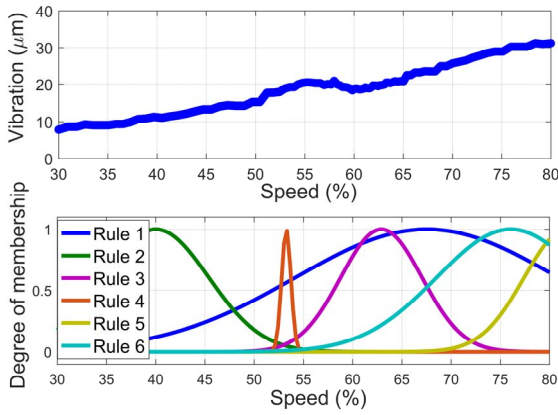


Fig. 6. Associated membership functions of the input: speed

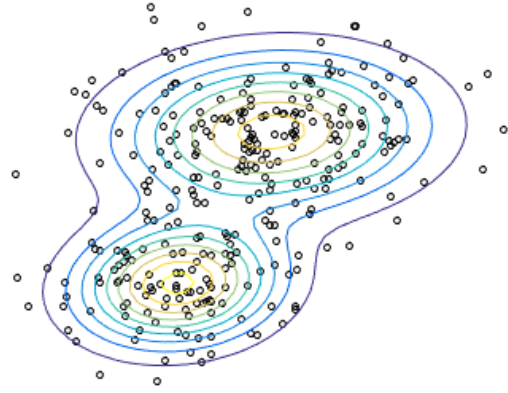


Fig. 7. Scatter plot and fitted GMM contour [18]

- *Decision making:* The most straightforward means of online decision making is to set thresholds on the health indices. The thresholds can be set by statistical analysis on the health indices to provide early warning on novelty and potential fault conditions, with the help of expert knowledge e.g. to identify 'normal' and 'abnormal' conditions. Thereafter, more testing sets are necessary for statistical analysis on threshold setting.

The other proposed method is based on Gaussian Mixture Model (GMM). The obtained (s, v) positions from training datasets are currently saved to form a (s, v) plane/cloud, where a Gaussian ellipse can be drawn, with confidence levels (e.g. 80%, 95%, etc.). It is expected that the identified range testing cases falling into will provide the 'confidence' of the start-up faultiness. Moreover, with the use of GMMs, different core configurations or sister engines can be separated by identifying the 'mixture' of Gaussians. This concept is demonstrated in Fig. 7, which is an example of GMM from Matlab® [18]. Again, this will require more training start-up datasets available for further analysis of the (s, v) plane, which will be considered in the future work.

#### IV. CONCLUSION

This paper has proposed a novelty detection strategy for IGT systems through the use of start-up vibration signature analysis. Here, ANFIS is used to extract the start-up vibration signature from 'healthy' start-up data through an offline training process. Then, for on-line testing, real start-up measurements are compared with the extracted signature. By matching the critical speed points, two health indices are introduced: s-health index is to show the difference on critical speed, and v-health index is to show the difference on vibration level at the critical speed. Comparing with conventional similarity measures, it is shown that the introduced s- and v-health indices can provide more insights in engine start-up health condition. Finally, experimental trials from running IGTs are studied to demonstrate the efficacy of the proposed approach. Future work in terms of industrial implementations is also discussed.

#### ACKNOWLEDGMENT

The authors would like to thank Siemens Industrial Turbomachinery, Lincoln, U.K., for providing research support and access to real-time data to support the research outcomes.

#### REFERENCES



- [1] Y. Zhang, C. Bingham, M. Gallimore, D. Cox, "Novelty detection based on extensions of GMMs for industrial gas turbines." *Proc. IEEE Int. Conf. on Computational Intelligence and Virtual Environments for Measurement Systems and Applications*, Shenzhen, China, Jun. 2015.
- [2] N. Baxter, "Forty-eight case histories of intriguing machinery problems." *Sound & Vibration*, vol. 46, no. 1, pp.36-51, 2012.
- [3] Y. Zhang, C. Bingham, Z. Yang, W.K. Ling, M. Gallimore, "Machine fault detection by signal denoising—with application to industrial gas turbines." *Measurement*, vol. 58, pp. 230-240, 2014.
- [4] S.K. Acharya, A. Das, V.K. Gupta, K.R. Anilkumar, Ravindranath, S. Bhattacharjee, "Vibration signature analysis as a diagnostic tool for condition assessment of rotating equipment – experience at TAPS." *Proc. of the National Seminar & Exhibition on Non-Destructive Evaluation*, NDE, Dec. 2009.
- [5] M. Seera, C.P. Lim, S. Nahavandi, C.K. Loo, "Condition monitoring of induction motors: A review and an application of an ensemble of hybrid intelligent models." *Expert Systems with Applications*, vol. 41, no. 10, pp. 4891-4903, 2014.
- [6] H.R. Depold, F.D. Gass, "The application of expert systems and neural networks to gas turbine prognostics and diagnostics," *J. of Engineering for Gas Turbines and Power*, vol. 121, no. 4, pp. 607-612, 1999.
- [7] R.B. Joly, S.O.T. Ogaji, R. Singh, S.D. Probert, "Gas-turbine diagnostics using artificial neural-networks for a high bypass ratio military turbofan engine," *Applied Energy*, vol. 78, no. 4, pp. 397-418, 2004.
- [8] Y. Zhang, J. Chen, C. Bingham, M. Mahfouf, "A new adaptive Mamdani-type fuzzy modeling strategy for industrial gas turbines," *Proc. of IEEE Int. Conf. on Fuzzy Systems*, pp. 1599 – 1603, Beijing, China, Jul. 2014.
- [9] J.S.R. Jang, "ANFIS: adaptive-network-based fuzzy inference system," *IEEE Transactions on Systems, Man and Cybernetics*, vol. 23, no. 3, pp. 665-685, 1993.
- [10] K. Salahshoor, M.S. Khoshro, M. Kordestani, "Fault detection and diagnosis of an industrial steam turbine using a distributed configuration of adaptive neuro-fuzzy inference systems," *Simulation Modelling Practice and Theory*, vol. 19, no. 5, pp. 1280-1293, 2011.
- [11] B. Shahriari, A.N. Shahrabaki, A. Shahriari, "Gas turbine fault detection and isolation using adaptive neurofuzzy inference system (ANFIS)," *Advanced Materials Research*, vol. 1016, pp. 721-725, 2014.
- [12] R.A. Saeed, A.N. Galybin, V. Popov, "3D fluid–structure modelling and vibration analysis for fault diagnosis of Francis turbine using multiple ANN and multiple ANFIS," *Mechanical Systems and Signal Processing*, vol. 34, no. 1–2, pp. 259–276, 2013.
- [13] S.H. Cha, "Comprehensive survey on distance/similarity measures between probability density functions," *Int. J. of Mathematical Models and Methods in Applied Science*, vol. 4, no. 1, pp. 300-307, 2007.
- [14] R.Q. Quiroga, A. Kraskov, T. Kreuz, P. Grassberger, "Performance of different synchronization measures in real data: a case study on electroencephalographic signals," *Phys. Review E*, vol. 65, 041903, 2002.
- [15] E. A. Ogbonnaya, E. M. Adigio, H. U. Ugwu, M. C. Anumiri, "Advance gas turbine rotor shaft fault diagnosis using Artificial Neural Network", *Int. J. of Engineering and Technology Innovation*, vol. 3, no. 1, pp. 58-69, 2013.
- [16] P. Tahmasebi, "A hybrid neural networks-fuzzy logic-genetic algorithm for grade estimation," *Computers & Geosciences*, vol. 42, pp. 18–27, 2012.
- [17] MATLAB<sup>®</sup>2014a, The MathWorks, Inc., Natick, Massachusetts, United States.
- [18] "Cluster Data from Mixture of Gaussian Distributions," MATLAB<sup>®</sup>2016a, The MathWorks, Inc., Natick, Massachusetts, United States.